Genes of *Bacillus cereus* and *Bacillus anthracis* Encoding Proteins of the Exosporium

Sarah J. Todd, Arthur J. G. Moir, Matt J. Johnson, and Anne Moir*

Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield, S10 2TN, United Kingdom

The exosporium is the outermost layer of spores of *Bacillus cereus* and its close relatives *Bacillus anthracis* and *Bacillus thuringiensis*. For these pathogens, it represents the surface layer that makes initial contact with the host. To date, only the BclA glycoprotein has been described as a component of the exosporium; this paper defines 10 more tightly associated proteins from the exosporium of *B. cereus* ATCC 10876, identified by N-terminal sequencing of proteins from purified, washed exosporium. Likely coding sequences were identified from the incomplete genome sequence of *B. anthracis* or *B. cereus* ATCC 14579, and the precise corresponding sequence from *B. cereus* ATCC 10876 was defined by PCR and sequencing. Eight genes encode likely structural components (*exsB*, *exsC*, *exsD*, *exsE*, *exsF*, *exsG*, *exsJ*, and *cotE*). Several proteins of the exosporium are related to morphogenetic and outer spore coat proteins of *B. subtilis*, but most do not have homologues in *B. subtilis*. ExsE is processed from a larger precursor, and the CotE homologue appears to have been C-terminally truncated. ExsJ contains a domain of GXX collagen-like repeats, like the BclA exosporium protein of *B. anthracis*. Although most of the exosporium genes are scattered on the genome, *bclA* and *exsF* are clustered in a region flanking the rhamnose biosynthesis operon; rhamnose is part of the sugar moiety of spore glycoproteins. Two enzymes, alanine racemase and nucleoside hydrolase, are tightly adsorbed to the exosporium layer; they could metabolize small molecule germinants and may reduce the sensitivity of spores to these, limiting premature germination.

Spores of the *Bacillus cereus* family, which includes *Bacillus anthracis* and *Bacillus thuringiensis*, all possess a loose balloon-like exosporium (7). A similar layer is also found on spores of some other bacilli and clostridia. The particular adherence and hydrophobic properties conferred by the exosporium (4, 14) suggest that it may possibly be of significance to spore pathogenicity. *Bacillus subtilis*, the paradigm of sporeformers, has no such clearly defined exosporial layer, so the exosporium has not been studied in molecular detail. Scanning electron microscopy has revealed a paracrystalline basal layer, with hexagonal periodicity, and a hairlike outer layer (3, 11). There are also pilus-like structures on the surface (15). The exosporium contains protein, lipid, and carbohydrate (43 to 52, 15 to 18, and 23% of dry weight, respectively [3, 19]). A spore glycoprotein of *B. thuringiensis* was purified and partially characterized (10); it was present as two forms according to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)—a 70-kDa monomer and a 205-kDa multimer. Another glycoprotein, BclA, important to the surface hairlike layer, has recently been identified in the *B. anthracis* exosporium (22). Crude exosporium extracts of *B. cereus* (5) contain at least 12 major and some minor protein components that can be readily solubilized, including GroEL and a zinc metalloprotease called immune inhibitor A, which is a recognized virulence factor in *B. thuringiensis* spores (8). These soluble proteins are unlikely to represent true integral structural components of the exosporium but are probably adsorbed onto the exosporium and may provide functional significance to this outer spore layer (5). We now identify proteins that resist salt and detergent extraction and are therefore candidate structural proteins of the exosporium.

MATERIALS AND METHODS

Spore preparation. Spores of *B. cereus* ATCC 10876 were prepared in CCY medium as described earlier (6), with incubation until the cultures contained in excess of 95% free spores. Spores were harvested by centrifugation (15,000 × g, 10 min) and were washed approximately eight times in ice-cold sterile distilled water, removing contaminating vegetative cells and debris from the surface of the pellet. Microscopic examination was used to determine if the spore suspension was free of vegetative cells and debris; if not, washing was continued. Pellets were resuspended in 50 mM Tris-HCl–0.5 mM EDTA (pH 7.5) and were stored at −20°C until further use.

Exosporium removal and purification. Spores (approximately 40 to 60 ml at 50 mg [dry weight] ml^{-1}) were passed twice through a French pressure cell at 21,700 lb/in^2. Exosporium fragments were separated from intact spores by pelleting the latter by low-speed centrifugation (9,000 × g for 5 min). The spore pellets were washed once in 50 mM Tris-HCl–0.5 mM EDTA (pH 7.5), and the exosporium-containing supernatants were pooled and then centrifuged again (10,000 × g, 15 min) to pellet any remaining spores.

The supernatant fractions containing exosporium fragments were concentrated by using an Amicon ultrafiltration cell fitted with a polyethersulfone membrane (nominal molecular mass cutoff, 10 kDa; Millipore). The concentrate was adjusted to contain 20% (vol/vol) Urografin 370 (Schering AG, Berlin, Germany), and samples of up to 2 ml were carefully layered onto 10 ml of 50% (vol/vol) Urografin 370 and were then centrifuged (14,000 × g for 30 min at 4°C); the exosporium-containing layer (the top 4 ml) was collected. Other material, such as cell debris and spores, pelleted at the bottom of the tube. The exosporium material was then dialyzed against at least four changes of distilled water at 4°C to eliminate Urografin 370 before centrifugation to pellet the exosporium fragments (40,000 × g, 90 min at 4°C). The exosporium pellet was resuspended in 50 mM Tris-HCl–0.5 mM EDTA (pH 7.5).

Salt and detergent washing of exosporium. The series of washes used was adapted from a method (20) for purifying *B. subtilis* spore coat fractions. Wash 1 was with TEP buffer (50 mM Tris-HCl, pH 7.2, 10 mM EDTA, and 2 mM

* Corresponding author. Mailing address: Department of Molecular Biology and Biotechnology, University of Sheffield, Firth Ct., Western Bank, Sheffield, S10 2TN, United Kingdom. Phone: 44 114 222 4418. Fax: 44 114 272 8697. E-mail: a.moir@sheffield.ac.uk.
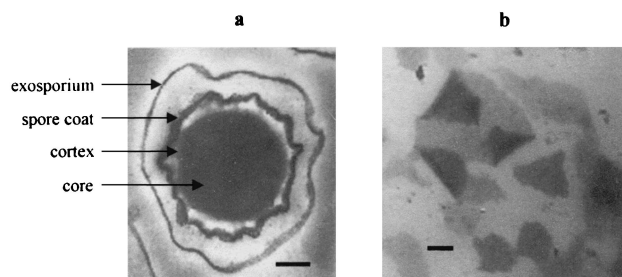
FIG. 1. Transmission electron micrographs of *B. cereus* ATCC 10876. These show negatively stained samples of a section of a spore (bar = 0.2 μm) (a) and of purified exosporium preparation (bar = 0.2 μm) (b).

phenylmethylsulfonyl fluoride) containing 0.5 M KCl and 1% (wt/vol) glycerol. Wash 2 was with 1 M NaCl. Wash 3 was with TEP buffer containing 0.1% SDS. Wash 4, to remove the SDS, was in TEP buffer alone. Wash 5 was with deionized water containing 0.01% (wt/vol) Tween 80, 2 mM phenylmethylsulfonyl fluoride, and 5 mM EDTA. All washes involved resuspension of purified exosporium pellets to 30 ml, followed by ultracentrifugation (184,000 × *g* for 1 h at 4°C) to pellet the exosporium. Between each set of washes, a small amount of the pellet was retained and resuspended in a small volume of 50 mM Tris-HCl–0.5 mM EDTA (pH 7.2) for analysis; the final resuspension was also made in the same buffer.

Electron microscopy. Spore sections were prepared and viewed as described in reference 17.

Exosporium samples were placed on Formvar-coated grids and were examined after negative staining by using a Philips CM10 transmission electron microscope at an accelerating voltage of 80 kV. Staining was with 1% (wt/vol) phosphotungstic acid (pH 7.2); after 15 to 30 s the excess phosphotungstic acid was withdrawn by using filter paper.

Gel electrophoresis. Samples were boiled for 4 min in sample buffer (50 mM Tris-HCl, pH 6.8; 10% glycerol; 2% SDS; 5% β-mercaptoethanol; and 0.1% bromophenol blue) and were centrifuged briefly (13,000 × *g*) to remove any insoluble material, and the proteins were separated by SDS-PAGE on 10% gels (16). Gels were stained with SYPRO Ruby Protein Gel Stain (Molecular Probes), silver stain (Bio-Rad), or Coomassie blue, as appropriate. Sigma protein molecular mass standards were used. To separate very small proteins, samples were boiled in different sample buffer (3 M Tris-HCl, pH 8.45; 12% glycerol; 4% SDS; and 0.1% Coomassie blue G) and were separated on precast 16% Tris-Tricine gels (Novex) by using the running buffer as recommended by the manufacturers. Novex Mark 12 protein standards were used, and the gels were stained with Coomassie blue. Protein concentrations were determined by the method described in reference 21.

N-terminal sequencing and glycoprotein staining. Proteins were electrophoretically transferred onto polyvinyl difluoride (Bio-Rad) and nitrocellulose (Amersham) membranes for N-terminal sequencing and glycoprotein staining, respectively, by using 10 mM CAPS [3-(cyclohexylamino)-1-propanesulfonic acid] transfer buffer (pH 11) containing 10% methanol. For proteins from Tris-Tricine gels, methanol was omitted from the transfer buffer. N-terminal sequences were determined by using automated Edman degradation in an Applied Biosystems gas phase sequencer. Glycoprotein staining was performed and visualized with a glycoprotein enhanced chemiluminescence detection kit (Amersham).

DNA sequencing. Sequences flanking exosporium genes in *B. anthracis* were used to design PCR primers to attempt PCR on the related *B. cereus* ATCC 10876 genomic DNA. Fragments of the expected sizes were amplified with Expand High Fidelity *Taq* polymerase (Roche) in the presence of 2.5 or 3 mM MgCl₂ and sequenced.

Nucleotide sequence accession number. The sequences have been submitted to GenBank as AY121972 (*exsB*), AF529877 (*exsC*), AY121974 (*exsD*), AF539613 (*exsE*), AY171090 (*exsF*), AY121975 (*exsG*), and AY183116 (*exsJ*).

RESULTS

Exosporium preparation. Figure 1a shows the position of the exosporium in the spore. Large volumes of spores were required to obtain a reasonable amount of exosporium; washed

exosporium preparations from 4 g (dry weight) of spores would contain only 2 to 6 mg of protein. Passing spores twice through the French press at 21,700 lb/in² caused partial fragmentation of the exosporium; samples of purified exosporium fragments (Fig. 1b) contained no other detectable particulate contaminating material, such as fragments of the spore coat. Coat protein profiles were very different from the exosporium profiles obtained (unpublished data), providing further evidence for the absence of contaminating materials in the exosporium preparation.

Protein profiles of exosporium fractions. Salt- and detergent-washed purified exosporium preparations yield a reproducible pattern of proteins, as shown by SDS-PAGE, which is very different from that of the unwashed preparations (Fig. 2). Adsorbed protein bands disappear, and the remaining bands correspond to very tightly bound or integral proteins of the exosporium. Salt washes removed immune inhibitor A (the prominent 73-kDa band) and most of the others, with the exception of bands at 46 and 21 kDa that were greatly reduced in the 0.1% SDS wash (wash 3). Newly enriched bands were seen (for example the 43-kDa band), and the 205-kDa band became relatively stronger. To separate small proteins (<25 kDa) that were not resolved on the gel in Fig. 2, exosporium samples were loaded onto 16% acrylamide, Tris-Tricine-buffered gels and were separated by SDS-PAGE (Fig. 3). Protein assays showed that approximately 75% of protein in unwashed, compared to 66% of protein in washed, exosporium preparations was solubilized with electrophoresis sample buffer.

Silver-stained gels. Silver staining was also used to compare the protein profiles of unwashed and washed exosporium (Fig. 4a). The overall profiles of bands, in washed and unwashed samples, were similar to those seen by Coomassie or SYPRO Ruby staining, but some bands became more prominent. In particular, a strong and diffuse band is seen from 70 kDa upwards (with a more intense region at 205 kDa).

Glycoprotein staining of standard SDS-PAGE gels showed the presence of carbohydrates in bands with molecular masses of approximately >205, 70, and 30 kDa (Fig. 4b). There was
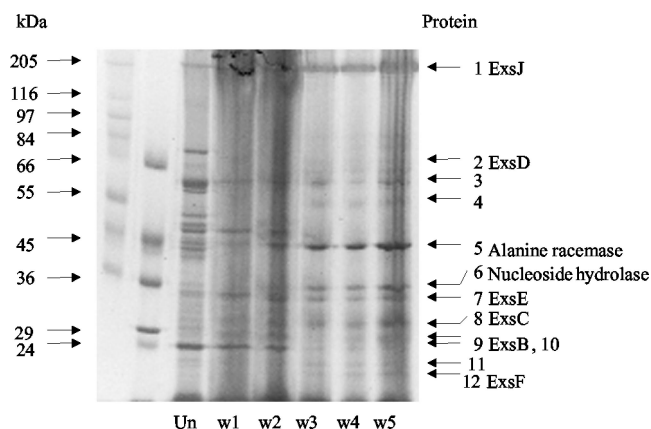

FIG. 2. Protein profile of unwashed exosporium compared to salt- and detergent-washed exosporium. Lanes show exosporium samples as follows: unwashed (Un) and pellets from wash 1 (w1), wash 2 (w2), wash 3 (w3), wash 4 (w4), and wash 5 (w5). Gel was 10% polyacrylamide and stained with SYPRO Ruby Protein Gel Stain. Washes were adapted from reference 20.
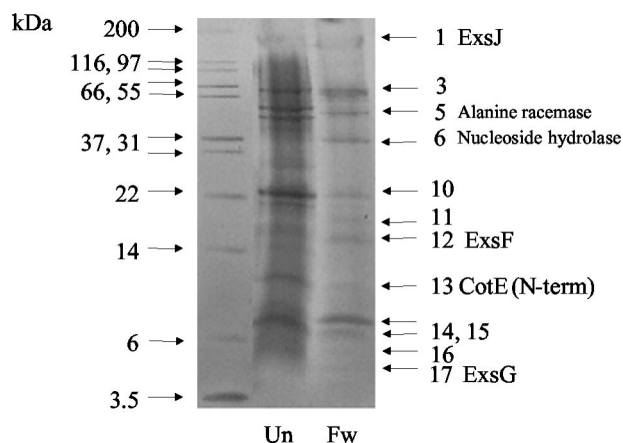
FIG. 3. Protein profile of smaller exosporium proteins comparing unwashed (Un) exosporium to fully washed (Fw) exosporium. Proteins were separated on precast 16% Tris-Tricine gels (Novex) by using the running buffer recommended by the manufacturers. Gel was stained with Coomassie blue.

TABLE 1. *B. cereus* ATCC 10876 exosporium protein properties (calculated with EMBOSS program at http://www.ebi.ac.uk) and N termini determined from gels

Gene name	Apparent molecular mass (kDa)	Predicted molecular mass (kDa)	Predicted pI	Sequenced N terminus
exsB	26	21[a]	10.0[a]	EDFLHQDP
exsC	30	16	4.7	THIIDYQATQPISKTGETT
exsD	66	18	5.2	ADYFYKDGKKYYKNQSH
exsE	34	22[a]	6.2[a]	KIVVDAGHQP
exsF	13	17	4.1	MFSSDCEFTKIDCEAKP
exsG	5	5	4.5	LLVTCILQEGNAF
exsJ	205, 70	42	4.6	MKHNDCFXHNNCNPIVF
cotE	12	20[b]	4.2[b]	SEFREIIXKAVVG

[a] Based on processed protein.
[b] Based on *B. cereus* ATCC 14579 CotE.

also staining material at the base of the stacking gel of too high a molecular weight to enter the gel. The >205- and 70-kDa bands correspond to regions that were strongly silver stained (Fig. 4a).

Identification of exosporium proteins. N-terminal data for protein bands were used to search preliminary sequence data for *B. anthracis* obtained from The Institute for Genomic Research website at http://www.tigr.org, by using tblastn (1), and to search open reading frames (ORFs) identified in preliminary data for *B. cereus* ATCC 14579 from http://www.integratedgenomics.com.

Identification of two enzymes in the exosporium. Band 5 protein (43 kDa) was identified from its N terminus (MEEA PFYRDTWVEVDLDAIYN) as an alanine racemase; as encoded in *B. anthracis*, it has 50% identity to the *B. subtilis dal* gene product, and its gene is located in the same genetic context between *ydcC* and *ydcD* homologues. A second alanine racemase homologue in the unfinished *B. anthracis* genome

FIG. 4. Glycoprotein staining of proteins from unwashed (Un) and fully washed (Fw) exosporium. Silver-stained 10% protein gel (a) is compared to blot (b), where glycoproteins have been visualized by using the enhanced chemiluminescence glycoprotein detection kit from Amersham. Arrows on the right indicate protein bands stained as glycoproteins.

sequence has a very different N-terminal amino acid sequence (only 10 out of 21 identities with the observed N terminus).

Band 6 protein (35 kDa; N terminus NKKIIFFGDFGIDD AVALI) corresponds precisely to that of an ORF in *B. anthracis* that is a predicted member of the inosine-uridine preferring nucleoside hydrolase family; this ORF shares conserved motifs and 30% amino acid identity with the paradigm protozoal enzyme (12).

Novel proteins identified from N-terminal sequencing. Other genes identified from these searches are novel; they are not present in the genome of *B. subtilis*, suggesting that many may be unique to the exosporium. These have therefore been given *exs* gene designations. Protein properties and experimental N termini are summarized in Table 1, and their complete sequences are in Fig. 5.

ExsB. Band 9 of Fig. 2 (molecular mass = 26.5 kDa) has an N terminus that is probably an internal sequence (from codon 18) of a single ORF in *B. anthracis*; although it is preceded by a Met residue, a likelier translational start with a good potential RBS is 18 residues earlier. If N-terminally processed, the presequence does not resemble a signal peptide. The mature 192-amino-acid *B. cereus* ExsB spore protein contains 20 cysteine residues and several short repeat sequences. The genomic context of *exsB* in both *B. cereus* and *B. anthracis* is equivalent to that of *cotG* in *B. subtilis*, as it is divergently transcribed from a *cotH* homologue (Fig. 6a). The *B. anthracis* CotH protein shows 55% amino acid identity to *B. subtilis* CotH. The *exsB* gene appears monocistronic, as a convergently transcribed phosphoglycerate mutase gene lies downstream (Fig. 6a).

Regions close to the N and C termini of ExsB have some similarity in composition to the equivalent regions in *B. subtilis* CotG, which also contains charged repeats and has a similar predicted pI of 10.3. The *B. cereus* ExsB protein shows 88% identity to the *B. anthracis* ExsB protein with the major difference at residues 17 to 34 of the mature *B. cereus* protein, where only 4 out of 18 residues are identical.

ExsC. Band 8 of Fig. 2 (apparent molecular mass of ca. 30 kDa) has an N-terminal sequence corresponding to an ORF (*exsC*) detectable in both *B. anthracis* and *B. cereus* DNA sequence databases. The 126-amino-acid protein is relatively hydrophilic and has a very limited similarity (28% identical over 104 residues to a sequence within the FliC flagellin of *Serratia marcescens* (SP_P13713). The *B. cereus* ATCC 10876
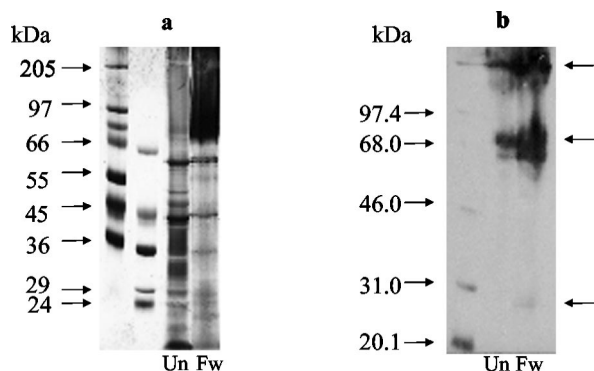
Protein	Protein sequence
ExsB	MKRDIRKAVEEIKSAGMEDFLHQDPSTFECDDDKFTHHHCTTGCKCTTGGKCPRTRCTRV KHCTFVTKCTHVKKWTFVTKCTRVRVQKWTFVTKVTRRKECVLVTKRTRRKHCTFITKCI RFEKKFFWTKRSFCKKCEFFPNRHGGSCDDSCDHGKDCHDSGHKWNDCKGGHKFPSCKNK KFDHFWYKKRNC
ExsC	MTHIIDYQATQPISKTGETTFAIPSSPNKAILANLKLRISSRDSRNNRVELIATIGIEGI TETSQVLFRIFRDNIEIFNAQVGIESTDSEQFYVQTFQAIDQNVSSGTHEYSLTVENLTS GASAEVVGPLSFSALAIGQERKCC
ExsD	MADYFYKDGKKYYKNQSHSNDQKNNCFIETHTIAGSAENENGNIPVSVFLETTAPQTVFE DFTNNHNKTLIQLFVVGMSAPVQVTILTRRSSVPITTTLQPVQTKIFQVEDFQSLTLTKQ EGSTSVVSLFVQKTFCICCKDNNDSCDEYYHECN
ExsE	MRTWRVGTFSMGLSIISLGCFLLFSVVKGIQVLDTLTAWWPVLLIILGAEVLLYLLFSKK EQSFIKYDIFSIFFIGVLGSVGIAFYCLLSTGLLEEVRHSINTTRQTSNIPDGQFDIPES IKKIVVDAGHQPLTIEGNNTNQIHLLGTYEMTTKANEKLNLKRDDFLSVQTAGETMYITL KSLPVQHTLFNSAPQVKPTLVLPQNKNVEIRASNNELSLYPGQLQNNWFVQESSRVSVHL AKESDVSLTAVTNQKETHGSTPWEQVEDLTKNENTSSEEHPELNTQEHWYKNSIKTGNGT YKLNIEKAYNLNMSVLEK
ExsF	MFSSDCEFTKIDCEAKPASTLPAFGFAFNASAPQFASLFTPLLLPSVSPNPNITVPVIND TVSVGDGIRILRAGIYQISYTLTISLDNVPTAPEAGRFFLSLNTPANIIPGSGTAVRSNV IGTGEVDVSSGVILINLNPGDLIQIVPVELIGTVDIRAAALTVAQIS
ExsG	MEFQLLVTCILQEGNAFFLVTKVDDVITLKVPITAGVAGLFLALGVPRCS
ExsJ	MKHNDCFDHNNCNPIVFSADCCKNPQSVPITREQLSQLITLLNSLVSAISAFFANPSNAN RLVLLDLFNQFLIFLNSLLPSPEVNFLKQLTQSIIVLLQSPAPNLGQLSTLLQQFYSALA QFFFALDLIPISCNSNVDSATLQLLFNLLIQLINATPGATGPTGPTGPTGPTGPAGTGAG PTGATGATGATGPTGATGPAGTGGATGATGVTGATGATGATGPTGPTGATGPTGATG ATGATGPTGATGPTGATGLTGATGAAGGGAIIPFASGTTPSALVNALVANTGTLLGFGFS QPGVALTGGTSITLALGVGDYAFVAPRAGTITSLAGFFSATAALAPISPVQVQIQILTAP AASNTFTVQGAPLLLTPAFAAIAIGSTASGIIAEAIPVAAGDKILLYVSLTAASPIAAVA GFVSAGINIV

FIG. 5. Exosporium protein sequences of *B. cereus* ATCC 10876. Underlined regions represent the N-terminal sequences obtained from protein bands on polyvinyl difluoride blots.

encoded protein is 96% identical to that of *B. cereus* ATCC 14579 but is much less conserved (66% identity) in *B. anthracis*, where the ORF has a very inefficient ATA initiation codon and may not be expressed. The ExsC protein of *B. cereus* migrates at the size predicted for a dimer.

ExsD. Band 2 of Fig. 2 appears as a weakly staining band of approximately 66 kDa and corresponds to an ORF of 154 amino acids, with five cysteine residues near its C terminus, in *B. cereus* and *B. anthracis*; these ExsD proteins show 94% identity. The *exsD* gene appears monocistronic. A second paralogous coding sequence in *B. anthracis* with an identical N terminus contains an internal stop codon. ExsD migrates very anomalously, or as a multimer, by SDS-PAGE.

ExsE. The N terminus shown in band 7 of Fig. 2 (apparent molecular mass = 34 kDa), corresponds to an internal sequence in a 318-amino-acid ORF, named *exsE*; this ORF appears to be translationally coupled to an upstream gene (*orf1*); thus, this hypothetical operon would be bicistronic (Fig. 6b). Two ORFs further upstream, one of which encodes a homologue of a *B. anthracis* ECF sigma 70 factor, could be part of the hypothetical operon. ExsE protein is predicted to contain an N-terminal 122-residue hydrophobic domain, with potentially three membrane spans, that is absent from the mature protein; the protein therefore would appear to be processed during assembly. The C-terminal hydrophilic domain, representing the assembled protein, comprises 196 amino acids and is relatively acidic. The mature ExsE protein shows 94% identity to the *B. anthracis* homologue.

ExsF. Band 12 in Fig. 2 shows an apparent molecular mass of 13 kDa. The *B. anthracis* genome contains an *exsF* gene (ExsF proteins are 97% conserved between *B. cereus* and *B. anthracis*) and also a distantly located paralogue; the latter

would encode a different N terminus that was not detected in our analysis. ExsF is acidic, and the 16% proline-plus-glycine content suggests that it may have a somewhat extended structure. As discussed later, the *exsF* gene is clustered with several other genes implicated in exosporium formation.

ExsG. Band 17 in Fig. 3 shows the smallest protein identified (apparent molecular mass = 5.4 kDa). Like ExsF, it is relatively acidic. It is encoded in a small monocistronic ORF whose predicted 50-residue sequence is identical in *B. cereus* and *B. anthracis*.

A processed CotE spore morphogenetic protein is located in the exosporium. The N-terminal sequence of a 12-kDa band

FIG. 6. Gene organization of the *exsB* (a) and *exsE* (b) regions in *B. cereus*. (a) The monocistronic *exsB* gene is situated between two convergently transcribed genes—homologues of *B. subtilis cotH* and phosphoglycerate mutase (*pgm*). (b) The *exsE* gene is situated in a hypothetical bicistronic operon as it appears to be translationally coupled to an upstream gene (*orf1*). Further upstream is a gene encoding a homologue of a *B. anthracis* ECF sigma 70 factor; a convergently transcribed gene downstream of *exsE* is a homologue of the *E. coli yjiN* gene.
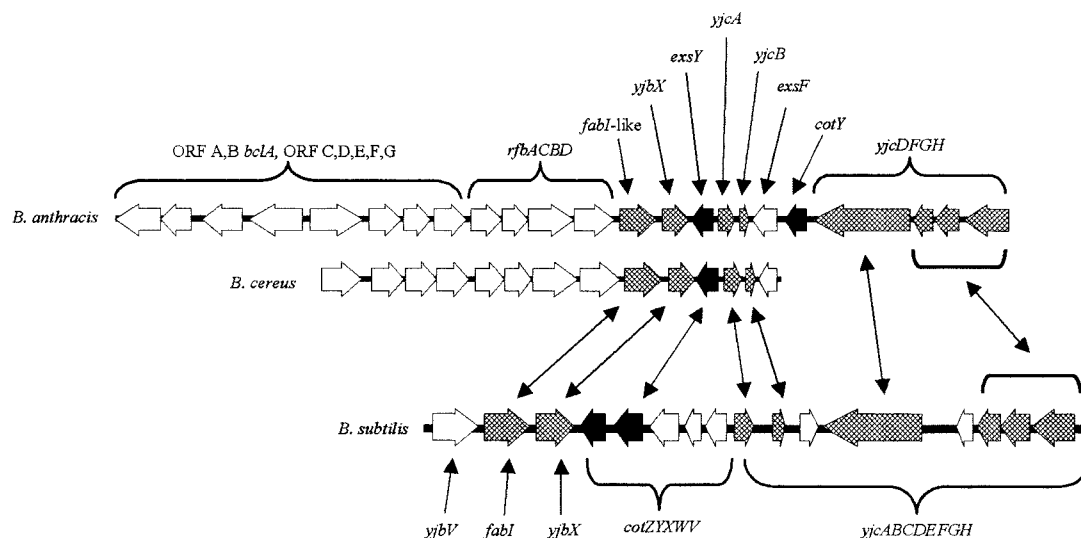
FIG. 7. Gene organization of the *exsF/bclA* region in *B. anthracis*, *B. cereus*, and *B. subtilis* showing local gene conservation. *B. cereus* data represent contig 1517 of the strain ATCC 14579 sequence. Cross-hatched genes denote genes conserved between *B. subtilis* and *B. anthracis*, white genes are nonconserved, and black genes are homologues of *cotY* and -*Z*. Gene names in clusters are ordered by their position in the genomes.

(Fig. 3, band 13) corresponded exactly to the N terminus (after the initial Met) of ORFs in *B. anthracis* and *B. cereus* ATCC 14579 that are homologues of *B. subtilis* CotE (59% identity). As the full-sized ORF encodes 180 amino acids, this protein band represents an N-terminal segment of the CotE protein. CotE is a major morphogenetic protein in spore coat assembly in *B. subtilis* and is situated between the inner and outer coat layers of the mature spore (18); deletion of its gene prevents the assembly of a major subset of coat proteins.

Glycoproteins in the *B. cereus* exosporium. A glycoprotein has been purified from the *B. thuringiensis* exosporium (10). We determined the N-terminal sequence of this purified protein, kindly provided by M. Garcia-Patrone, as MKHNDXF (where X indicates uncertainty at that position), and it was identical for both 205-kDa and monomeric 70-kDa forms. In *B. cereus*, the protein shown in band 1 of Fig. 2 (205-kDa) stains heavily as a glycoprotein (Fig. 4b); its N-terminal sequence was determined to be MKHNDCFXHNNCNPIVF. The similarity in size and identical N-terminal sequence suggests that this protein corresponds to the Garcia-Patrone protein purified from the *B. thuringiensis* exosporium. This protein does not correspond to the BclA glycoprotein, although we have independent evidence of the presence of the latter in the *B. cereus* exosporium (C. Redmond, A. Moir, and K. Bailey-Smith, unpublished data). No individual ORF in the unfinished *B. cereus* ATCC 14579 or in the *B. anthracis* genome sequences corresponded exactly to the band 1 N terminus. The *B. cereus* ATCC 14579 sequence near the paralogue of ExsF, however, contains two adjacent ORFs (ExsI and ExsH) that both contain elements similar to this sequence, but neither has the identical N terminus. Both have a GPX-type collagen-like repeat also found in the recently characterized BclA glycoprotein from *B. anthracis* (22). PCR has confirmed the gene arrangement for the *exsH* and -*I* region in *B. cereus* ATCC 10876 as identical to that reported for the unfinished *B. cereus* ATCC 14579. Primers that flank *exsH* did not, however, amplify the equivalent *exsH* gene from *B. thuringiensis* subsp. *kurstaki*

DNA; instead, a gene was amplified (named *exsJ*) whose product has the N terminus identical to that of our our N-terminally sequenced *B. cereus* and *B. thuringiensis* glycoprotein. The *exsJ* gene of *B. cereus* 10876 has since been identified by PCR and sequenced; the ExsJ protein is indeed a very close homologue of ExsH (81% identity). The main differences are in their N-terminal domains (70% identity) that precede the collagen-like region; the C-terminal domains are almost identical. Subsequent DNA sequencing (S. J. Todd, unpublished data) has shown that the *B. thuringiensis* subsp. *kurstaki* ExsJ protein is identical to that of *B. cereus* ATCC 10876.

The broad silver-stained glycoprotein band at 70 kDa has been seen only with these silver-stained gels and with glycoprotein staining; its N-terminal sequence has not been obtained but is likely to be monomeric ExsJ (as shown for *B. thuringiensis*). A fainter glycoprotein band at 30 kDa is also of unknown provenance. Any attempt to relate them to identified Exs proteins will have to await individual gene inactivation.

DISCUSSION

The solubilizable material (66% of preparation) from salt- and detergent-washed exosporium contains few recognized spore coat protein homologues of *B. subtilis*—only the morphogenetic protein CotE. From the 17 distinguishable bands representing suspected structural or tightly adsorbed components of the exosporium, we have identified 10 gene products. Many of these are entirely novel (ExsB to ExsG) and show unusual amino acid composition and/or possible posttranslational modification. They are generally not clustered in the genome of *B. anthracis*.

The *exsF* gene (though not its paralogue) is located in a gene cluster (Fig. 7) near the *exsY* gene, which is a *cotY* homologue required for exosporium assembly (M. J. Johnson, Todd, and Moir, unpublished data); there is also a second homologue of the *cotY* gene that is transcribed convergently from *exsF*. This

cluster is separated from the *bclA* region by the rhamnose biosynthesis operon (*rfbACBD*); the *bclA* gene is clustered with glycosyltransferases and methyltransferases (Fig. 7). Rhamnose and methyl rhamnose are major carbohydrate groups of the exosporium (9, 24) that are present in the characterized glycoprotein of *B. thuringiensis* (10). At least the majority of genes in this region are therefore likely to be implicated in exosporium formation. There is also local conservation between this region in *B. cereus* and *B. anthracis* and a region in *B. subtilis*. The *cotVWXYZ* gene cluster in *B. subtilis* is flanked on one side by *fabI/yjbX* and on the other by *yjcA* to *-H*. A similar organization is seen in *B. cereus* ATCC 14579 and *B. anthracis*, although the *exsY/exsF/cotY* region replaces the *cot* gene cluster and the *yjc* gene cluster lacks *yjcC* and *yjcE* (Fig. 7). PCR and sequencing have confirmed the *exsY/exsF/cotY* arrangement in *B. cereus* ATCC 10876.

From the available unfinished genome sequences, most of the novel Ex proteins are closely conserved between *B. cereus* and *B. anthracis*, with two exceptions discussed earlier—a local region of ExsB and the entire ExsC protein that may not be expressed in *B. anthracis*. Differences between *exs* genes may be useful for comparative studies or for specific detection of *B. cereus* versus *B. anthracis*.

Some proteins may not be structural elements but may be merely very strongly adsorbed. This is certainly the case for two enzymes, alanine racemase and nucleoside hydrolase (bands 5 and 6, respectively [Fig. 2]). Two major chemical triggers of spore germination in *B. cereus* are alanine and inosine (2, 6, 23). Alanine racemase converts L-alanine to D-alanine, a competitive inhibitor of germination, while a nucleoside hydrolase would degrade inosine. Inclusion of an alanine racemase inhibitor increases the germination rate of *B. cereus* spores (13). It is likely that both of these enzymes moderate the spore's germination rate.

The presence of CotE suggests a link between coat and exosporium assembly in *B. cereus*; transposon mutagenesis (Todd and Moir, unpublished) confirms this. The protein composition of the exosporium is clearly complex, and this specialized layer has some elements in common with, and some very distinct from, the *B. subtilis* spore coat. The glycoprotein complement of the exosporium is also likely to be complex, as there are multiple related genes.

The identified genes do not by any means represent an exhaustive list of protein components of the exosporium; one-third of protein remained in the insoluble fraction, and 7 out of 17 bands have not yielded clear N-terminal sequence data.

There is much more to learn of the role of the novel proteins in assembly or structure; site-specific mutations are presently being constructed and the effects on exosporium formation are being examined.

ADDENDUM IN PROOF

The *B. anthracis* homologue of exosporium protein ExsF has recently been described and designated BxpB (C. Steichen, P. Chen, J. F. Kearney, and C. L. Turnbough, Jr., J. Bacteriol. **185:**1903–1910, 2003).

REFERENCES

1. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST—a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
2. **Barlass, P. J., C. W. Houston, M. O. Clements, and A. Moir.** 2002. Germination of *Bacillus cereus* spores in response to L-alanine and to inosine: the roles of *gerL* and *gerQ* operons. Microbiology **148:**2089–2095.
3. **Beaman, T. C., H. S. Pankratz, and P. Gerhardt.** 1971. Paracrystalline sheets reaggregated from solubilized exosporium of *Bacillus cereus*. J. Bacteriol. **107:**320–324.
4. **Bowen, W. R., A. S. Fenton, R. W. Lovitt, and C. J. Wright.** 2002. The measurement of *Bacillus mycoides* spore adhesion using atomic force microscopy, simple counting methods, and a spinning disc technique. Biotechnol. Bioeng. **79:**170–179.
5. **Charlton, S., A. J. G. Moir, L. Baillie, and A. Moir.** 1999. Characterization of the exosporium of *Bacillus cereus*. J. Appl. Microbiol. **87:**241–245.
6. **Clements, M. O., and A. Moir.** 1998. The role of the *gerI* operon of *Bacillus cereus* 569 in the response of spores to germinants. J. Bacteriol. **180:**6729–6735.
7. **Desrosier, J. P., and J. C. Lara.** 1984. Synthesis of the exosporium during sporulation of *Bacillus cereus*. J. Gen. Microbiol. **130:**935–940.
8. **Fedhila, S., P. Nel, and D. Lereclus.** 2002. The InhA2 metalloprotease of *Bacillus thuringiensis* strain 407 is required for pathogenicity in insects infected via the oral route. J. Bacteriol. **184:**3296–3304.
9. **Fox, A., G. E. Black, K. Fox, and S. Rostovtseva.** 1993. Determination of carbohydrate profiles of *Bacillus anthracis* and *Bacillus cereus* including identification of O-methyl methylpentoses by using gas chromatography-mass spectrometry. J. Clin. Microbiol. **31:**887–894.
10. **Garcia-Patrone, M., and J. S. Tandecarz.** 1995. A glycoprotein multimer from *Bacillus thuringiensis* sporangia: dissociation into subunits and sugar composition. Mol. Cell. Biochem. **145:**29–37.
11. **Gerhardt, P., and E. Ribi.** 1964. Ultrastructure of the exosporium enveloping spores of *Bacillus cereus*. J. Bacteriol. **88:**1774–1789.
12. **Gopaul, D. N., S. L. Meyer, M. Degano, J. C. Sacchettini, and V. L. Schramm.** 1996. Inosine-uridine nucleoside hydrolase from *Crithidia* fasciculate. Genetic characterization, crystallization, and identification of histidine 241 as a catalytic site residue. Biochemistry **35:**5963–5970.
13. **Gould, G. W.** 1966. Stimulation of L-alanine-induced germination of *Bacillus cereus* spores by D-cycloserine and O-carbamyl-D-serine. J. Bacteriol. **92:**1261–1262.
14. **Koshikawa, T., M. Yamazaki, M. Yoshimi, S. Ogawa, A. Yamada, K. Watabe, and M. Torii.** 1989. Surface hydrophobicity of spores of Bacillus species. J. Gen. Microbiol. **135:**2717–2722.
15. **Kozuka, S., and K. Tochikubo.** 1985. Properties and origin of filamentous appendages on spores of *Bacillus cereus*. Microbiol. Immunol. **29:**21–37.
16. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680–685.
17. **Leatherbarrow, A. J. H., M. A. Yazdi, J. P. Curson, and A. Moir.** 1998. The *gerC* locus of *Bacillus subtilis*, required for menaquinone biosynthesis, is concerned only indirectly with spore germination. Microbiology **144:**2125–2130.
18. **Little, S., and A. Driks.** 2001. Functional analysis of the *Bacillus subtilis* morphogenetic spore coat protein CotE. Mol. Microbiol. **42:**1107–1120.
19. **Matz, L. L., T. C. Beaman, and P. Gerhardt.** 1970. Chemical composition of exosporium from spores of *Bacillus cereus*. J. Bacteriol. **101:**196–201.
20. **Nicholson, W. L., P. Setlow, A. Galizzi, and R. Sammons.** 1990. Sporulation, germination and outgrowth, p. 391–450. *In* C. R. Harwood and S. M. Cutting (ed.), Molecular biological methods for bacillus. John Wiley & Sons, Chichester, United Kingdom.
21. **Peterson, G. L.** 1977. A simplification of the protein assay method of Lowry which is more generally applicable. Anal. Biochem. **83:**346–356.
22. **Sylvestre, P., E. Couture-Tosi, and M. Mock.** 2002. A collagen-like surface glycoprotein is a structural component of the *Bacillus anthracis* exosporium. Mol. Microbiol. **45:**169–178.
23. **Warren, S. C., and G. W. Gould.** 1968. *Bacillus cereus* spore germination: absolute requirement for an amino acid. Biochim. Biophys. Acta **170:**341–350.
24. **Wunschel, D., K. F. Fox, G. E. Black, and A. Fox.** 1995. Discrimination among the *Bacillus cereus* group, in comparison to *Bacillus subtilis*, by structural carbohydrate profiles and ribosomal RNA spacer region PCR. Syst. Appl. Microbiol. **17:**625–635.